



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/638,082	08/14/2000	Jeffrey A. Dean	Google-3 (GOOGP008)	1030
44989	7590	10/26/2006	EXAMINER	
HARRITY SNYDER, LLP 11350 Random Hills Road SUITE 600 FAIRFAX, VA 22030			BASHORE, WILLIAM L	
			ART UNIT	PAPER NUMBER
			2176	

DATE MAILED: 10/26/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/638,082	DEAN ET AL.	
	Examiner	Art Unit	
	William L. Bashore	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 25 August 2006.

2a) This action is **FINAL**. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-3,5,7-14,16-24 and 26 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1-3,5,7-14,16-24 and 26 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
 3) Information Disclosure Statement(s) (PTO/SB/08)
 Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
 Paper No(s)/Mail Date. _____

5) Notice of Informal Patent Application
 6) Other: _____

DETAILED ACTION

1. This action is responsive to communications: amendment filed 8/25/2006, to the original application filed 8/14/2000, with provisional filing date 4/6/2000.
2. It is noted that the previous indication of allowable subject matter has been withdrawn in view of newly found art. Accordingly, this action is non-final.
3. Claims 1-3, 5, 7-14, 16-24, 26 pending. Claims 1, 10, 12, 20, 22, 23, 26 are independent claims.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. **Claims 1-3, 5, 7-14, 16-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Najork et al. (hereinafter "Najork '265"), US 6,321,265 B1 filed 11/2/1999 in view of Hoffert et al. (hereinafter Hoffert), US 6,374,260 B1 filed 2/28/2000.**

Regarding independent claim 1, Najork '265 teaches sending a request for additional links to hyperlinked documents to a link manager in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as a link manager for the web crawler. The Frontier data structure provides links to the web crawler. Najork '265 teaches receiving a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches grouping the

plurality of links to hyperlinked documents by host in fig. 7 and col. 2 lines 24-36. Najork '265

teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60

- col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches crawling a

hyperlinked document from the selected host in col. 1 lines 31-47 and col. 3 lines 3-52.

Najork '265 does not teach that the plurality of links to be crawled are selected by the link manager based on priority. However, Hoffert teaches that the plurality of links (pages) to be crawled are selected based on priority (Hoffert col. 2 lines 30-36, col. 4 line 12-18). It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Hoffert into Najork '265 to have created the claimed invention. It would have been obvious and desirable to have implemented a prioritized selection of the links so that media rich content or rapidly changing would have higher priority and crawled first, facilitating quicker indexing (Hoffert column 4 lines 15-17).

Regarding dependent claim 2, Najork '265 teaches wherein the stall time of the host is the earliest time in which a hyperlinked document from the host should be crawled in fig. 6 col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Regarding dependent claim 3, Najork '265 teaches selecting a host with a stall time that is earlier than the current time in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Regarding dependent claim 5, Najork '265 teaches examining hyperlinked documents to be crawled at each host until a host is found with a stall time that is earlier than the current time in fig. 5-7, col. 1 line 31 - col. 2 line 2, and col. 2 lines 37-62. Najork '265 does not teach examining the groups in the specific "descending" order of the number of hyperlinked documents to be crawled. It would have been obvious to one of ordinary skill in the art at the time of the

invention to have modified Najork '265 to have taught this limitation. It would have been obvious and desirable to have examined the hosts in a descending order of hyperlinked documents to be crawled so that the hosts were examined in a logical order. This would have made the link selection process for the web crawler straightforward and easy to program.

Regarding dependent claim 7, Najork '265 does not teach moving the selected host to a group with one less hyperlinked documents to be crawled. It would have been obvious to one of ordinary skill in the art at the time the invention was made to have modified Najork '265 so that the a selected host would have been moved to a group with one less hyperlinked documents to be crawled. It would have been obvious and desirable to have done this so that the groups would have remained balanced for the web crawler.

Regarding dependent claim 8, Najork '265 teaches determining a retrieval time for retrieving the hyperlinked document from the selected host in col. 2 lines 43-52.

Regarding dependent claim 9, Najork '265 teaches adjusting subsequent stall times for the selected host according to the retrieval times in col. 2 lines 43-52.

Regarding independent claim 10, Najork '265 teaches computer code that requests links from a link manager in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as a link manager for the web crawler. The Frontier data structure provides links to the web crawler. Najork '265 teaches receiving a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches grouping the plurality of links to hyperlinked

Art Unit: 2176

documents by host in fig. 7 and col. 2 lines 24-36. Najork '265 teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches a computer readable medium that stores computer codes in fig. 1 and col. 1 lines 13-30.

Najork '265 does not teach that the plurality of links to be crawled are selected by the link manager based on priority. However, Hoffert teaches that the plurality of links (pages) to be crawled are selected based on priority (Hoffert col. 2 lines 30-36, col. 4 line 12-18). It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Hoffert into Najork '265 to have created the claimed invention. It would have been obvious and desirable to have implemented a prioritized selection of the links so that media rich content or rapidly changing would have higher priority and crawled first, facilitating quicker indexing (Hoffert column 4 lines 15-17).

Regarding dependent claim 11, Najork '265 teaches a computer readable medium which is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive in fig. 1 and col. 1 lines 13-30.

Regarding independent claim 12, Najork '265 teaches sending a request for additional links to hyperlinked documents to a link manager in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as a link manager for the web crawler. The Frontier data structure provides links to the web crawler. Najork '265 teaches receiving a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches grouping the plurality of links to hyperlinked documents by host in fig. 7 and col. 2 lines 24-36. Najork '265 teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2

line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches determining a retrieval time for retrieving the hyperlinked document from the selected host in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches adjusting subsequent stall times for the selected host according to the retrieval time in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Najork '265 does not teach that the plurality of links to be crawled are selected by the link manager based on priority. However, Hoffert teaches that the plurality of links (pages) to be crawled are selected based on priority (Hoffert col. 2 lines 30-36, col. 4 line 12-18). It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Hoffert into Najork '265 to have created the claimed invention. It would have been obvious and desirable to have implemented a prioritized selection of the links so that media rich content or rapidly changing would have higher priority and crawled first, facilitating quicker indexing (Hoffert column 4 lines 15-17).

Regarding dependent claim 13, Najork '265 teaches wherein the stall time of the host is the earliest time in which a hyperlinked document from the host should be crawled in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Regarding dependent claim 14, Najork '265 teaches selecting a host with a stall time that is earlier than the current time in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Regarding dependent claim 16, Najork '265 teaches examining the groups in descending order of the number of hyperlinked documents to be crawled at each host until a host is found with a stall time that is earlier than the current time in fig. 5-7, col. 1 line 31 - col. 2 line 2, and col. 2 lines 37-62.

Regarding dependent claim 17, Najork '265 teaches sorting the hosts in fig. 6-7 and col. 1 line 60 col. 2 line 2 and col. 2 lines 37-62. Najork '265 does not specifically teach sorting the hosts by stall time. It would have been obvious to one of ordinary skill in the art at the time the invention was made to have modified Najork '265 to have sorted the hosts according to stall time. It would have been obvious and desirable to have done this so that the web crawler could have crawled the sites in a time-efficient order.

Regarding dependent claim 18, Najork '265 does not teach moving the selected host to a group with one less hyperlinked documents to be crawled. It would have been obvious to one of ordinary skill in the art at the time the invention was made to have modified Najork '265 so that the a selected host would have been moved to a group with one less hyperlinked documents to be crawled. It would have been obvious and desirable to have done this so that the groups would have remained balanced for the web crawler.

Regarding dependent claim 19, Najork '265 teaches determining a retrieval time for retrieving the hyperlinked document from a selected host in col. 3 lines 23-29.

Regarding independent claim 20, Najork '265 teaches sending a request for additional links to hyperlinked documents to a link manager in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as a link manager for the web crawler. The Frontier data structure provides links to the web crawler. Najork '265 teaches receiving a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches grouping the plurality of links to hyperlinked documents by host in fig. 7 and col. 2 lines 24-36.

Najork '265 teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches determining a retrieval time for retrieving the hyperlinked document from the selected host in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches adjusting subsequent stall times for the selected host according to the retrieval time in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches a computer readable medium that stores computer codes in fig. 1 and col. 1 lines 13-30.

Najork '265 does not teach that the plurality of links to be crawled are selected by the link manager based on priority. However, Hoffert teaches that the plurality of links (pages) to be crawled are selected based on priority (Hoffert col. 2 lines 30-36, col. 4 line 12-18). It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Hoffert into Najork '265 to have created the claimed invention. It would have been obvious and desirable to have implemented a prioritized selection of the links so that media rich content or rapidly changing would have higher priority and crawled first, facilitating quicker indexing (Hoffert column 4 lines 15-17).

Regarding dependent claim 21, Najork '265 teaches a computer readable medium which is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive in fig. 1 and col. 1 lines 13-30.

6. **Claims 22-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Najork et al. (hereinafter "Najork '265"), US 6,321,265 B1 filed 11/2/1999 in view of Najork et al. (hereinafter "Najork '755"), US 6,351,755 B1 filed 11/2/1999.**

Regarding independent claim 22, Najork '265 teaches storing a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47. Najork '265 teaches receiving additional links to hyperlinked documents in fig. 1 and col. 3 lines 3-52. Najork '265 teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2 line 2 and col. 2 lines 37-62. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47.

Najork '265 teaches determining that more links to hyperlinked documents are desired and sending requests to multiple link managers for more links to hyperlinked documents in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as link managers for the web crawler. Each link queue manages links from a specific host to provide to the web crawler and thus is similar to each of the claimed link managers.

Najork '265 does not explicitly teach sending requests to multiple link managers for more links to hyperlinked documents. However, Najork '755 teaches queues containing URL's from multiple server hosts, as well as teaching multiple queues implemented in various scenarios (Najork '755 column 16 lines 10-26). it would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Najork '755 into Najork '265 and modified the Frontier queuing and indexing system of the combination to have operated the queues as individual link managers so that links could have still be provided to the web crawler in the event one of the queues experienced an interrupted connection with the web crawler.

Regarding independent claim 23, Najork '265 teaches storing a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47. Najork '265 teaches receiving additional links to hyperlinked documents in fig. 1 and col. 3 lines 3-52. Najork '265 teaches

selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2 line 2 and col. 2 lines 37-62. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47. Najork '265 teaches a computer readable medium that stores computer codes in fig. 1 and col. 1 lines 13-30.

Najork '265 teaches determining that more links to hyperlinked documents are desired and sending requests to multiple link managers for more links to hyperlinked documents in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as link managers for the web crawler. Each link queue manages links from a specific host to provide to the web crawler and thus is similar to each of the claimed link managers.

Najork '265 does not explicitly teach sending requests to multiple link managers for more links to hyperlinked documents. However, Najork '755 teaches queues containing URL's from multiple server hosts, as well as teaching multiple queues implemented in various scenarios (Najork '755 column 16 lines 10-26). It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Najork '755 into Najork '265 and modified the Frontier queuing and indexing system of the combination to have operated the queues as individual link managers so that links could have still been provided to the web crawler in the event one of the queues experienced an interrupted connection with the web crawler.

Regarding dependent claim 24, Najork '265 teaches a computer readable medium which is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive in fig. 1, and col.. 1 lines 13-30.

7. **Claim 26 is rejected under 35 U.S.C. 103(a) as being unpatentable over Najork '265.**

Regarding independent claim 26, Najork '265 teaches sending a request for additional links to hyperlinked documents to a link manager in fig. 2-4 and col. 5 line 53 - col. 6 line 6. The Frontier data structure queues organize and manage the links to hyperlinked documents and act as a link manager for the web crawler. The Frontier data structure provides links to the web crawler. Najork '265 teaches receiving a plurality of links to hyperlinked documents to be crawled in col. 1 lines 31-47 and col. 3 lines 3-52. Najork '265 teaches grouping the plurality of links to hyperlinked documents by host in fig. 7 and col. 2 lines 24-36. Najork '265 teaches selecting a host to crawl next according to a stall time of the host in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39. Najork '265 teaches crawling a hyperlinked document from the selected host in col. 1 lines 31-47 and col. 3 lines 3-52.

Najork '265 teaches selecting a host with a stall time that is earlier than the current time in fig. 6, col. 1 line 60 - col. 2 line 2, col. 2 lines 37-62, and col. 3 lines 23-39.

Najork '265 does not teach examining the groups in the specific "descending" order of the number of hyperlinked documents to be crawled. However, it would have been obvious to one of ordinary skill in the art at the time of the invention to have modified Najork '265 to have taught this limitation. It would have been obvious and desirable to have examined the hosts in an order (such as descending order) of hyperlinked documents to be crawled so that the hosts were examined in a logical order. This would have made the link selection process for the web crawler straightforward and easy to program.

Response to Arguments

8. Applicant's arguments with respect to the pending claims have been considered but are moot in view of the new grounds of rejection.

Conclusion

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to William L. Bashore whose telephone number is (571) 272-4088. The examiner can normally be reached on 11:30am - 8:00pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on (571) 272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

10. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

William L. Bashore
WILLIAM BASHORE
PRIMARY EXAMINER

October 24, 2006